

Algebraic Ordinals*

S. L. Bloom[†]

Department of Computer Science
Stevens Institute of Technology
Hoboken, NJ, USA

Z. Ésik[‡]

Department of Informatics
University of Szeged
Szeged, Hungary

February 8, 2010

Abstract

An **algebraic tree** T is one determined by a finite system of fixed point equations. The frontier $\text{Fr}(T)$ of an algebraic tree T is linearly ordered by the lexicographic order $<_\ell$. If $(\text{Fr}(T) <_\ell)$ is well-ordered, its order type is an **algebraic ordinal**. We prove that the algebraic ordinals are exactly the ordinals less than ω^{ω^ω} .

1 Introduction

Fixed points and finite systems of fixed point equations occur in just about all areas of computer science. Regular and context-free languages, rational and algebraic formal power series, finite state process behaviors can all be characterized as (components of) canonical solutions (e.g., unique, least or greatest, or initial solutions) of systems of fixed point equations.

Consider the fixed point equation

$$X = 1 + X$$

over linear orders, where $+$ denotes the sum operation (functor) on linear orders. As explained in [BE10], its canonical (initial) solution is the ordinal ω , or any

*This is a revised version of the paper submitted July 5, 2009

[†]Supported in part by the TAMOP-4.2.2/08/1/2008-0008 program of the Hungarian National Development Agency.

[‡]Supported in part by grant no. K 75249 from the National Foundation of Hungary for Scientific Research and by the TAMOP-4.2.2/08/1/2008-0008 program of the Hungarian National Development Agency.

linear order isomorphic to the ordering of the natural numbers. For another example, consider the system of fixed point equations

$$\begin{aligned} X &= Y + X \\ Y &= 1 + Y \end{aligned}$$

The first component of its canonical solution is ω^2 , and the second component is ω . Of course, there exist fixed point equations whose canonical solution is not well-ordered, for example, the canonical solution of

$$X = X + 1$$

is ω^* , the reverse of ω , and the canonical solution of

$$X = X + 1 + X$$

is the ordered set of the rationals.

The above equations are quite simple since they involve no parameters. The unknowns X, Y range over linear orders, or equivalently, constant functions (or rather, functors) defined on linear orders. By allowing unknowns ranging over functions (or functors) in several variables, we obtain the ordinal ω^ω as the first component of the canonical solution of

$$\begin{aligned} F_0 &= G(1) \\ G(x) &= x + G(F(x)) \\ F(x) &= x + F(x) \end{aligned}$$

The second and third components of the canonical solution are functors which map a linear order x to $x \times \omega^2$ and $x \times \omega$, respectively.

We call a linear order algebraic if it is isomorphic to the first (or principal) component of the canonical solution of a system of fixed point equations of the sort

$$F_i(x_0, \dots, x_{n_i-1}) = t_i, \quad i = 1, \dots, n \tag{1}$$

where $n_1 = 0$ and each t_i is an expression composed of the function variables F_j , $j = 1, \dots, n$, the individual variables x_0, \dots, x_{n_i-1} , the constant 1 and the sum operation $+$. Moreover, we call a linear order regular if it is isomorphic to the first component of the canonical solution of a system (1) with $n_i = 0$ for all i . Further, we call an ordinal algebraic or regular if it is an algebraic or regular linear order.

From the results in [Heil80], it follows easily that an ordinal is regular if and only if it is less than ω^ω (See also [BC01].) It was conjectured in [BE07] that an ordinal is algebraic if and only if it is less than ω^{ω^ω} . Our aim in this paper is to confirm this conjecture.

Infinite structures may be described by finite presentations in several different ways. One method, represented by automatic structures, consists in describing a structure up to isomorphism by representing its elements as words or trees or some other combinatorial objects, and its relations by rewriting rules or automata. Another, algebraic approach is describing an infinite structure as the canonical solution of a system of fixed point equations over a suitably defined algebra of structures. Our results show that these methods are equivalent, at least for small ordinals.¹

An automatic relational structure [Hod82, KN95] is a countable structure whose carrier is given by a regular language and whose relations can be computed by synchronous multi-tape automata. A tree automatic structure [DT90, Del04, Col04] is defined using tree automata, cf. [GS84]. It was proved in [Del04] that the automatic ordinals are exactly the ordinals less than ω^ω . See also [KRS03]. In [Del04], it is also shown that an ordinal is tree automatic if and only if it is less than ω^{ω^ω} . Thus, an ordinal is automatic if and only if it is regular², and is tree automatic if and only if it is algebraic. Actually the claim that every regular ordinal is automatic is immediate. It would be interesting to derive a direct proof of the fact that every algebraic ordinal is tree automatic, or the other way around.

In a traditional setting, one solves a system of fixed point equations in an algebra equipped with a suitable partial order; there is a least element, suprema of ω -chains exist, the operations preserve the ordering and least upper bounds of ω -chains. Such algebras are commonly called continuous algebras (or ω -continuous algebras), cf. [GTWW77, Gue81]. In this setting, one solution of this kind of system is provided by least fixed points. The classical Mezei-Wright theorem [MW67] asserts that such a solution is preserved by a continuous, order preserving algebra homomorphism.

However, in several settings such as (countable) linear orders, there is no well-defined partial order but one can naturally introduce a category by considering morphisms between linear orders. A generalization of the classical Mezei-Wright theorem to the setting of “continuous categorical algebras” has been given in [BE10], where least elements are replaced by initial elements, and suprema of ω -chains are replaced by colimits of ω -diagrams. Since trees, equipped with the usual partial order [GTWW77, Gue81, BE10] form an initial continuous categorical algebra, it follows that instead of solving a system of fixed point equations directly over linear orders, we may first find its least solution over trees, to obtain an algebraic or regular tree, and then take the image of this tree with respect to the unique morphism from trees to linear orders which assigns to a tree the linear order determined by the frontier of the tree. Thus, up to isomorphism, an algebraic (or regular) linear order is the frontier of an algebraic (or regular) tree. In this way, we may represent algebraic and regular linear orders and ordinals as frontiers of algebraic or regular trees, and this is

¹For another result in this direction, see [Col04].

²Compare this fact with Theorem 2 in [BKN08].

the approach we take here. This approach is not new. Courcelle was the first to use frontiers of trees to represent linear orders and words, i.e., labeled linear orders. The origins of the notions of regular linear order and regular word go back to [Cour78a].

In our argument showing that every algebraic ordinal is less than $\omega^{\omega^{\omega}}$, we will make use of certain context-free grammars generating prefix languages, called “ordinal grammars” which seem to have independent interest. Equipped with the lexicographic order (see below), the language generated by an ordinal grammar is well-ordered. We show that an ordinal is the order type of a language generated by an ordinal grammar if and only if it is less than $\omega^{\omega^{\omega}}$. We then show how to translate an algebraic tree given by a system of equations to an ordinal grammar generating the frontier of the tree.

The paper is organized as follows. In Section 2, we define regular and algebraic trees and linear orders, and regular and algebraic ordinals. Then, in Section 3, we establish some closure properties of algebraic linear orders, including closure under sum, multiplication, and ω -power, and use these closure properties to establish that any ordinal less than $\omega^{\omega^{\omega}}$ is algebraic. Section 4 is devoted to ordinal grammars and the proof of the result that an ordinal is the order type of the context-free language generated by an ordinal grammar, equipped with the lexicographic order, if and only if it is less than $\omega^{\omega^{\omega}}$. Then, in Section 5 we show how to construct for an algebraic tree T (given by a system of fixed point equations) an ordinal grammar G such that the order type of the language generated by G equals the order type of the frontier of T . (The proof of the correctness of the translations is moved to an appendix.) The paper ends with Section 6 containing some concluding remarks.

2 Linear orders, words, prefix languages, and trees

2.1 Linear orders

A linearly ordered set $(A, <_A)$ consists of a set A equipped with a strict linear order, i.e., an irreflexive and transitive binary relation $<_A$ that satisfies exactly one of the conditions

$$x = y \text{ or } x <_A y \text{ or } y <_A x,$$

for all $x, y \in A$. If $(A, <_A)$ and $(B, <_B)$ are linearly ordered sets, a **morphism** $h : (A, <_A) \rightarrow (B, <_B)$ is a function $h : A \rightarrow B$ such that for all $x, y \in A$, if $x <_A y$ then $h(x) <_B h(y)$. The collection of all linearly ordered sets and morphisms forms a category **LO**. We say two linearly ordered sets are isomorphic if they are isomorphic in **LO**. The **order type** of a linear order is the isomorphism class of the linear order in **LO**. We write $\mathbf{o}(A, <_A)$ for the order type of $(A, <_A)$.

A linearly ordered set $(A, <_A)$ is **well-ordered** if every nonempty subset of A has a least element. If α is an ordinal, then α is well-ordered by the membership relation \in . When $(A, <_A)$ is well-ordered, there is a unique ordinal α such that (α, \in) is isomorphic to $(A, <_A)$, and the order type of $(A, <_A)$ is conveniently identified with this ordinal.

Remark. We remind the reader that if $\alpha = \mathbf{o}(A, <_A)$ and $\beta = \mathbf{o}(B, <_B)$ are ordinals, where A, B are disjoint sets, then $\alpha + \beta$ is the order type of $A \cup B$, ordered by putting every element of A before every element of B ; otherwise, imposing the given ordering on elements inside A or B . Similarly, if $\alpha_n = \mathbf{o}(A_n, <_n)$ is an ordinal for each $n \geq 0$, where the sets A_n are pairwise disjoint, then $\sum_n \alpha_n$ is the order type of the union $\bigcup_n A_n$ ordered by making every element of A_n less than every element in A_m , for $n < m$, and imposing the given order on A_n .

The ordinal $\alpha \times \beta$ is the order type of $A \times B$, ordered by “last differences”, i.e., where

$$(a, b) < (a', b') \iff b <_B b' \text{ or } (b = b' \text{ and } a <_A a').$$

Last, α^ω is the least upper bound of all ordinals $\alpha^n = \overbrace{\alpha \times \dots \times \alpha}^n$, for $n < \omega$.

For more definitions and facts on linear orders and ordinals we refer to [Roit90, Ros82, Sier58]. For later use, we prove:

Lemma 2.1 *Suppose that α is an ordinal and β is an infinite ordinal, so that $1 + \beta = \beta$. If α is a successor ordinal then*

$$(\beta + 1) \times \alpha \leq \beta \times \alpha + 1.$$

If α is 0 or a limit ordinal, then

$$(\beta + 1) \times \alpha \leq \beta \times \alpha.$$

Proof. We prove both claims by transfinite induction on α . The case when $\alpha = 0$ is clear. Assume that α is a successor ordinal. Then $\alpha = \gamma + n$ for some positive integer n where γ is 0 or a limit ordinal. Then, using the induction hypothesis and the assumption $1 + \beta = \beta$,

$$\begin{aligned} (\beta + 1) \times \alpha &= (\beta + 1) \times \gamma + (\beta + 1) \times n \\ &\leq \beta \times \gamma + \beta \times n + 1 \\ &= \beta \times (\gamma + n) + 1 \\ &= \beta \times \alpha + 1. \end{aligned}$$

Suppose now that α is a limit ordinal. Then,

$$\begin{aligned} (\beta + 1) \times \alpha &= \sup_{\gamma < \alpha} ((\beta + 1) \times \gamma) \\ &\leq \sup_{\gamma < \alpha} (\beta \times \gamma + 1) \\ &\leq \beta \times \alpha, \end{aligned}$$

again by the induction hypothesis. \square

2.2 Words

We use \mathbb{N} for the set of nonnegative integers. If $n \in \mathbb{N}$, then

$$[n] = \{0, \dots, n-1\},$$

so that $[0] = \emptyset$. The collection of all finite words on a set A is denoted A^* . The empty word is denoted ϵ . The set of nonempty words on A is $A^+ = A^* \setminus \{\epsilon\}$. We denote the product (concatenation) of two words $u, v \in A^*$ by uv . We identify an element $a \in A$ with the corresponding word of length one.

When A is linearly ordered by $<_A$, then A^* is equipped with two partial orders. The **prefix order**, written $u <_p v$, is defined by:

$$u <_p v \iff v = uw, \quad w \neq \epsilon,$$

for some word w . The **strict** or **branching order**, written $u <_s v$ is defined by: $u <_s v$ if and only if for some words u_1, u_2, v_2 , and $a, b \in A$,

$$\begin{aligned} u &= u_1 a u_2 \\ v &= u_1 b v_2, \quad \text{and} \\ a &<_A b. \end{aligned}$$

The **lexicographic order** on A^* is defined by:

$$u <_\ell v \iff u <_p v \text{ or } u <_s v.$$

It is clear that for any $u, v \in A^*$, exactly one of the following possibilities holds:

$$u = v, \quad u <_p v, \quad v <_p u, \quad u <_s v, \quad v <_s u.$$

Thus, the lexicographic order is a linear order on A^* .

2.3 Prefix languages

A **prefix language** on a set A is a subset L of A^* such that if $u \in L$ and $uv \in L$ then $v = \epsilon$. For later use, we state the following facts about prefix languages.

Lemma 2.2 • If $(P, <_P)$ is a countable linearly ordered set, there is a prefix language L on $[2]$ such that $(P, <_P)$ and $(L, <_\ell)$ are isomorphic.

- If L is a prefix language, then, for $u, v \in L$, $u <_\ell v \iff u <_s v$.
- If $K, L \subseteq A^*$ are prefix languages, where A is linearly ordered, so is $KL = \{uv : u \in K, v \in L\}$ and, when $(K, <_\ell)$ and $(L, <_\ell)$ are well-ordered,

$$\mathbf{o}(KL, <_\ell) = \mathbf{o}(L, <_\ell) \times \mathbf{o}(K, <_\ell).$$

- Suppose that $L \subseteq [2]^*$ is a prefix language such that $\alpha = \mathbf{o}(L, <_\ell)$ is an ordinal. Then

$$\left(\bigcup_{n \geq 0} 1^n 0 L^n, <_\ell \right)$$

is well-ordered and, if $\alpha > 1$,

$$\alpha^\omega = \mathbf{o}\left(\bigcup_{n \geq 0} 1^n 0 L^n, <_\ell\right). \quad (2)$$

Proof. For the first statement we refer to [Cour78a, BE04]. We prove only the statement (2). The set, say L_∞ , of words on the right side of (2) is the set of all words of the form $1^n 0 u$, for $u \in L^n$. For any words u, v , if $1^n 0 u <_p 1^m 0 v$, then $n = m$ and $u <_p v$. Since L is a prefix language, so is L^n and, thus so is $\bigcup_n 1^n 0 L^n$. Also, if $n < m$, $1^n 0 u <_s 1^m 0 v$, and $1^n 0 u <_s 1^n 0 v$ if and only if $u <_s v$. Thus, by definition, the order type of the lexicographic order of the prefix language L_∞ is $\sum_n \alpha^n$. It is now easy to see that this sum is α^ω . \square

2.4 Trees

Suppose that Σ is a (finite) ranked alphabet, i.e., a nonempty finite set partitioned into subsets Σ_k of “ k -ary operation symbols”, $k \geq 0$. Moreover, suppose that $V = \{x_0, x_1, \dots\}$ is an ordered set of “individual variables”. A Σ -**tree** T in the variables V , or a “tree over Σ in the individual variables V ”, is a partial function $[n]^* \rightarrow \Sigma \cup V$ satisfying the conditions listed below. Here, n is largest such that $\Sigma_n \neq \emptyset$. The conditions are:

- The domain of T , $\text{dom}(T)$, is prefix closed: if $T(uv)$ is defined, then so is $T(u)$.
- If $T(u) \in \Sigma_k$, $k > 0$, and $T(ui)$ is defined, then $i \in [k]$.
- If $T(u) \in \Sigma_0 \cup V$, then u is a leaf, and $T(ui)$ is undefined, for all i .

A Σ -tree T is **complete** if whenever $T(u)$ is defined in Σ_n , for some u , then $T(u0), \dots, T(u(n-1))$ are all defined. Moreover, T is **finite** if its domain is finite. Below we will usually denote finite trees by lower case letters. The size of a finite tree t is the size of the set $\text{dom}(t)$. The set of all Σ -trees in the variables V is denoted $T_\Sigma^\omega(V)$. Moreover, for a subset $V_n = \{x_0, \dots, x_{n-1}\}$ of V , we write $T_\Sigma^\omega(V_n)$ for the collection of all trees all whose leaves are labeled in $\Sigma_0 \cup V_n$. When $n = 0$, we write simply T_Σ^ω .

Trees are equipped with the following partial order $T \sqsubset T'$: Given $T, T' \in T_\Sigma^\omega(V)$ such that $T \neq T'$, we define $T \sqsubset T'$ if and only if for all words u , if $T(u)$ is defined, then $T(u) = T'(u)$. It is well-known that the partially ordered set $(T_\Sigma^\omega(V), \sqsubset)$ is ω -complete, i.e., $T_\Sigma^\omega(V)$ has as least element \perp , the totally undefined tree, and least upper bounds of all ω -chains. Indeed, if $T_0 \sqsubset T_1 \sqsubset \dots$ is an ω -chain in $T_\Sigma^\omega(V)$, the supremum is the tree T whose domain is the union of the domains of the trees T_n , and, if u is a word in this union, $T(u) = \sigma \in \Sigma$ if and only if $T_n(u) = \sigma$, for some n . Similarly, $T(u) = v_i$ for some $v_i \in V$ if and only if $T_n(u) = v_i$ for some n . Each symbol $\sigma \in \Sigma_k$ induces a k -ary operation on $T_\Sigma^\omega(V)$ in the usual way. It is well-known that these operations are ω -continuous in all arguments and $T_\Sigma^\omega(V)$ is an ω -continuous algebra. In the same way, for every n , $T_\Sigma^\omega(V_n)$ is an ω -continuous algebra (in fact, the free ω -continuous Σ -algebra on V_n). See [GTWW77, Gue81].

2.5 Tree substitution

In this section we define a substitution operation on trees, sometimes called second-order substitution.

Suppose that Σ and Δ are ranked alphabets and for each $\sigma \in \Sigma_n$ we are given a tree $R_\sigma \in T_\Delta^\omega(V_n)$. We define substitution in two steps, first for finite trees, and by continuity for infinite trees. For each finite tree $t \in T_\Sigma(V)$ we define the tree $S = t[\sigma \mapsto R_\sigma]_{\sigma \in \Sigma}$ in $T_\Delta^\omega(V)$, sometimes denoted just $t[\sigma \mapsto R_\sigma]$ by induction on the size of t . When t is the empty tree \perp , so is S . When t is x , for some $x \in V$, then $S = x$. Otherwise t is of the form $\sigma(t_0, \dots, t_{n-1})$, where $\sigma \in \Sigma_n$, and we define

$$S = R_\sigma(t'_1, \dots, t'_n)$$

where $t'_i = t_i[\sigma \mapsto R_\sigma]$, for all i .

Suppose now that T is an infinite tree in $T_\Sigma(V)$. Then there is an ascending ω -chain (t_n) of finite trees such that $T = \sup_n t_n$. We define

$$T[\sigma \mapsto R_\sigma] = \sup_n t_n[\sigma \mapsto R_\sigma].$$

It is known, see [Cour83], that substitution is ω -continuous.

Proposition 2.3 *Substitution is a continuous function*

$$T_{\Sigma}^{\omega}(V) \times \prod_n T_{\Delta}^{\omega}(V_n)^{\Sigma_n} \rightarrow T_{\Delta}^{\omega}(V).$$

Below when Σ and Δ are not disjoint and $R_{\sigma} = \sigma(x_1, \dots, x_n)$ for some $\sigma \in \Sigma_n$, then we often omit σ from the arguments of the substitution.

2.6 Algebraic trees and ordinals

Now consider a finite system E of equations of the form

$$\begin{aligned} F_1(x_0, \dots, x_{n_1-1}) &= t_1(x_0, \dots, x_{n_1-1}) \\ F_2(x_0, \dots, x_{n_2-1}) &= t_2(x_0, \dots, x_{n_2-1}) \\ &\vdots \\ F_m(x_0, \dots, x_{n_m-1}) &= t_m(x_0, \dots, x_{n_m-1}), \end{aligned} \tag{3}$$

where, for $i = 1, \dots, m$, t_i is a **term** over the ranked alphabet $\Sigma \cup \mathcal{F}$ in the variables $\{x_0, \dots, x_{n_i-1}\}$ (i.e., finite complete tree in $T_{\Sigma}^{\omega}(V_{n_i})$), where $\mathcal{F} = \{F_1, \dots, F_m\}$ is the set of “function variables” and each F_i has rank n_i . See the example in (4) below. Each term t_i induces a function

$$t_i^E : T_{\Sigma}^{\omega}(V_{n_1}) \times \dots \times T_{\Sigma}^{\omega}(V_{n_m}) \rightarrow T_{\Sigma}^{\omega}(V_{n_i})$$

by substitution:

$$(R_1, \dots, R_m) \mapsto t_i[F_j \mapsto R_j]_{1 \leq j \leq m}.^3$$

By Proposition 2.3 this function is ω -continuous. The target tupling

$$\langle t_1^E, \dots, t_m^E \rangle$$

mapping $T_{\Sigma}^{\omega}(V_{n_1}) \times \dots \times T_{\Sigma}^{\omega}(V_{n_m})$ to itself is also ω -continuous and has a least fixed point (T_1, \dots, T_m) . Thus, there is a least solution (T_1, \dots, T_m) of any such system. One function variable F_i of rank 0 is selected as the **principal variable**, and the corresponding tree T_i is the **principal component** of the least solution. (Typically, we choose the first function variable as the principal variable.) If every integer n_i , $i = 1, \dots, m$, is zero, the system is said to be **regular**.

Definition 2.4 *A tree T in $T_{\Sigma}^{\omega}(V_k)$ is **algebraic in** $T_{\Sigma}^{\omega}(V_k)$, (respectively, **regular**), if there is a finite system E , (respectively, **regular system**), of equations as above such that T is the principal component of the least solution of E in $T_{\Sigma}^{\omega}(V_k)$.*

³It is understood here that each $\sigma \in \Sigma_k$ remains unchanged, i.e., gets substituted by $\sigma(x_0, \dots, x_{k-1})$.

An alternative definition is possible by interpreting finite systems of fixed point equations directly on the “continuous categorical algebra” of linear orders. See [BE10].

We will use the above definition mainly when $k = 0$. It is known that when $T \in T_\Sigma^\omega(V_n)$ and $n < m$, then T is algebraic in $T_\Sigma^\omega(V_n)$ if and only if it is algebraic in $T_\Sigma(V_m)$. Thus, we may simply call T just algebraic. Moreover, we say that a tree $T \in T_\Sigma^\omega(V)$ is algebraic if it is algebraic in $T_\Sigma(V_n)$, for some $n \geq 0$. It is also known that when $\Sigma \subseteq \Sigma'$, then a tree $T \in T_\Sigma^\omega(V)$ is algebraic if and only if T is algebraic in $T_{\Sigma'}^\omega(V)$. Similar facts and conventions hold for regular trees. So below we can simply say that a tree is algebraic, or regular without specifying exactly the ranked alphabet.

The set of leaves of a tree T in $T_\Sigma^\omega(V)$ is denoted

$$\text{Fr}(T) = \{u \in [n]^* : T(u) \in \Sigma_0 \cup V\},$$

and is called the **frontier** of T . $\text{Fr}(T)$ is a prefix language linearly ordered by $<_\ell$. Here, n is the maximum of the ranks of the symbols in Σ .

Definition 2.5 *A linear order is **algebraic** (respectively, **regular**) if it is isomorphic to $(\text{Fr}(T), <_\ell)$ for some algebraic (respectively, regular) tree. An algebraic or regular ordinal is an ordinal which is an algebraic or regular linear order.*

Heilbrunner [Heil80] proved that the frontiers of regular trees are those obtainable from the empty and one point frontiers by the operations of “concatenation, omega and omega-op powers”, and infinitely many “shuffle” operations. It is an easy corollary of this fact that the regular ordinals are those less than ω^ω . (There is a somewhat longer argument based only on the facts in [BC01].) In this paper we will prove that an ordinal is algebraic if and only if it is less than ω^{ω^ω} .

The following fact is known. See [Cour83].

Proposition 2.6 *The classes of algebraic and regular trees are closed under first- and second-order substitution.*

In the remaining part of this section we show that for algebraic and regular linear orders one may restrict attention to those algebraic or regular Σ -trees for which $\Sigma_n = \emptyset$ unless $n = 2$ or $n = 0$.

Let Δ be the ranked alphabet with one binary function symbol and one constant symbol; otherwise, $\Delta_n = \emptyset$.

Proposition 2.7 *For any ranked alphabet Σ and any algebraic tree $T \in T_\Sigma^\omega(V)$ there is an algebraic tree $T' \in T_\Delta^\omega$ such that $(\text{Fr}(T), <_\ell)$ and $(\text{Fr}(T'), <_\ell)$ are isomorphic. \square*

For example, consider the system E of equations

$$\begin{aligned} F_0 &= \sigma_1(a, b, F_1(a)) \\ F_1(x) &= F_2(x, x) \\ F_2(x, y) &= \sigma_1(\sigma_2(a), F_2(x, F_2(x, y)), y) \end{aligned} \tag{4}$$

which uses a function symbol σ_1 in Σ_3 . The least solution consists of three trees (T_0, T_1, T_2) having vertices of out-degree 3. We replace the system E by the system

$$\begin{aligned} F_0 &= g(a, g(a, F_1(a))) \\ F_1(x) &= F_2(x, x) \\ F_2(x, y) &= g(a, g(F_2(x, F_2(x, y)), y)) \end{aligned}$$

in which the right hand terms use only the function variables and the one binary function symbol g , and the one constant symbol a . If (T'_0, T'_1, T'_2) is the least solution of this second system, $(\text{Fr}(T_i), <_\ell)$ is isomorphic to $(\text{Fr}(T'_i), <_\ell)$, for $i = 1, 2, 3$.

Thus, from now on, we will assume that if T is an algebraic tree, then $\text{Fr}(T) \subseteq [2]^*$.

Example 2.8 *Let Σ contain the binary symbol g , the unary symbol f and the constant a . Consider the system*

$$\begin{aligned} F_0 &= F(a) \\ F(x) &= g(x, F(f(x))) \end{aligned}$$

Then the first component of the least solution of this system is the tree

$$T_0 = g(a, g(f(a), g(f(f(a)), \dots, g(f^n(a), \dots))))).$$

Thus, this tree is algebraic. See also [Cour78b, Cour83, Gue81].

3 Closure properties of algebraic ordinals

In this section we use certain closure properties of algebraic ordinals to prove that every ordinal less than ω^{ω^ω} is algebraic.

Proposition 3.1 *Let \mathcal{C} be any set of ordinals which contains 0, 1, and is closed under sum, product and ω -power: i.e., if $\alpha, \beta \in \mathcal{C}$, then $\alpha + \beta$, $\alpha \times \beta$, α^ω belong to \mathcal{C} . Then all ordinals less than ω^{ω^ω} belong to \mathcal{C} .*

Proof. This follows from the assumptions and induction, making use of the Cantor Normal Form [Roit90, Ros82, Sier58] for ordinals less than ω^{ω^ω} . In fact, it is known that the set of ordinals less than ω^{ω^ω} is the least set of ordinals containing 0, 1 which is closed under sum, product, and ω -power. \square

We use Proposition 3.1 to show all ordinals less than ω^{ω^ω} are algebraic. For this reason, we fix the ranked alphabet Δ containing only a binary symbol g and a unary symbol a . We show that all ordinals less than ω^{ω^ω} arise as frontiers of algebraic trees in T_Δ^ω .

The ordinal 0 is algebraic, since if T is the empty tree, then $\text{Fr}(T)$ is the empty language, and the order type of the empty language is 0. The ordinal 1 is algebraic since the one-point tree is algebraic.

Suppose that $\alpha = \mathbf{o}(\text{Fr}(T), <_\ell)$ and $\beta = \mathbf{o}(\text{Fr}(S), <_\ell)$, where T and S are algebraic trees in T_Δ^ω .

Proposition 3.2 *If α and β are algebraic ordinals, so is $\alpha + \beta$.*

Proof. Consider the algebraic tree $g(T, S) = g(a, b)[a \mapsto T, b \mapsto S]$ whose root is labeled by the function symbol g and whose left subtree is T and whose right subtree is S . Then the tree $g(T, S)$ is algebraic, and its frontier has order type $\alpha + \beta$. \square

Since 0 and 1 are algebraic, we have the easy corollary that *every finite ordinal is algebraic*.

Proposition 3.3 *If α and β are algebraic ordinals, so is $\alpha \times \beta$.*

Proof. The tree $S[a \mapsto T]$ is algebraic, and its frontier has order type $\alpha \times \beta$. \square

Proposition 3.4 *If α is an algebraic ordinal, so is α^ω .*

Proof. Suppose that $\alpha = \mathbf{o}(\text{Fr}(T), <_\ell)$ for an algebraic tree T in T_Δ^ω . Consider the tree T_0 of Example 2.8 and let S be the algebraic tree in T_Δ^ω obtained by substituting the tree T for each vertex labeled f : $S = T_0[f \mapsto T]$. Then S is algebraic and its frontier is of order type α^ω . \square

Corollary 3.5 *Every ordinal less than ω^{ω^ω} is algebraic.*

Proof. By Propositions 3.2, 3.3 and 3.4, together with Proposition 3.1. \square

4 Grammars

In our argument proving that all algebraic ordinals are less than $\omega^{\omega^{\omega}}$ we will use certain context-free grammars, called ordinal grammars.

Throughout this section, we assume that

$$G = (N, T, S, P)$$

is a context-free grammar, with nonterminals N , start symbol $S \in N$, terminals $T = \{0, 1\}$, and productions P .

We will denote finite words on the alphabet $\{0, 1\}$ by u, v, w, x, y, \dots ; nonterminals will be written X, Y, Z, \dots , and we will denote by $p, q, r, s \dots$ words on $N \cup T$, possibly containing nonterminals.

Further, we assume that each context-free grammar has the following properties:

- Either each nonterminal $X \in N$ is “coaccessible”, i.e., $L(X)$ is a nonempty subset of $\{0, 1\}^*$, where $L(X)$ is the collection of all words $u \in \{0, 1\}^*$ such that there is some derivation

$$X \xRightarrow{*} u,$$

or $N = \{S\}$ and P is empty. We write $L(G)$ for $L(S)$.

- Each nonterminal X is “accessible”, i.e., there is some derivation

$$S \xRightarrow{*} qXr$$

where $q, r \in (N \cup T)^*$.

We end this section with some classical definitions. Suppose X, Y are nonterminals. Write

$$Y \preceq X$$

if there is some derivation $X \xRightarrow{*} pYq$ for some p and q . Define $X \approx Y$ if both $X \preceq Y$ and $Y \preceq X$ hold. When $Y \preceq X$ but $X \approx Y$ does not hold, we write $Y \prec X$.

The relation \preceq is a preorder on the nonterminals, and induces a partial order \leq on the equivalence classes

$$[X] := \{Y : X \approx Y\},$$

where $[Y] \leq [X]$ if $Y \preceq X$.

We say X is a **recursive nonterminal** if there is some *nontrivial derivation*

$$X \xRightarrow{*} pXq, \tag{5}$$

for some words $p, q \in (N \cup T)^*$. If not, X is a **non-recursive nonterminal**. When X is non-recursive, and $X \rightarrow p$ is a rule, $Y \prec X$, for all nonterminals in p .

Definition 4.1 The *height* of a nonterminal X is the number of equivalence classes $[Y]$ strictly below $[X]$. If q is a finite word on $N \cup T$, the height of q , $\mathbf{ht}(q)$, is the maximum of the heights of the nonterminals occurring in q . If q has no nonterminals, $\mathbf{ht}(q) = -1$.

If there are k nonterminals, $\mathbf{ht}(X) < k$, for all nonterminals X . If X has height zero, and $Y \preceq X$, then $Y \approx X$.

For any word q in $(N \cup T)^*$, write $L(q)$ for all words in $\{0, 1\}^*$ derivable in G from q .

4.1 Prefix and ordinal grammars

Our definition of an ordinal grammar is motivated by Proposition 4.3 below. But first we need the following fact:

Lemma 4.2 Suppose $L \subseteq \{0, 1\}^*$ is any language. If $(L, <_\ell)$ is not well-ordered, then there is a countable descending chain (u_n) , $n = 0, 1, \dots$, of words in L such that

$$u_{n+1} <_s u_n,$$

for each $n \geq 0$.

Proof. Suppose that (v_n) is a countable $<_\ell$ -descending chain of words in L . Then, for each n , either $v_{n+1} <_p v_n$ or $v_{n+1} <_s v_n$. Now define $u_0 = v_0$. Since v_0 has only finitely many prefixes, there is a least k such that $v_{k+1} <_s v_k <_p \dots <_p v_0$. Then $u_1 = v_{k+1} <_s u_0$, since $u <_s v$ if $u <_s w$ and $w <_p v$, for any words u, v, w . Similarly, assuming that u_m has been defined as $v_{m'}$, for some m' , we may define u_{m+1} as the first v_k such that $k > m'$ and $v_k <_s u_m$. \square

We take note of the following inheritance property.

Proposition 4.3 If $(L(G), <_\ell)$ is well-ordered, then, for any nonterminal X , $(L(X), <_\ell)$ is also well-ordered.

Proof. If not, by Lemma 4.2, suppose that there is a countable chain

$$\dots <_s u_1 <_s u_0.$$

of words in $L(X)$. Thus, for each n , $u_{n+1} = x_n 0 y_n$ and $u_n = x_n 1 z_n$, for some words $x_n, y_n, z_n \in \{0, 1\}^*$. Since all nonterminals are accessible and coaccessible, there are words v, w such that

$$vu_i w \in L(G),$$

for each $i \geq 0$, and thus

$$vu_{n+1} w <_s vu_n w,$$

for each $n \geq 0$, showing $(L(G), <_\ell)$ is not well-ordered, contradicting the hypothesis. \square

Definition 4.4 A grammar G is a **prefix grammar** if, for each nonterminal X , $L(X)$ is a prefix language. An **ordinal grammar** is a prefix grammar such that $(L(G), <_\ell)$ is well-ordered.

If $G = (N, T, P, S)$ is an ordinal grammar, then by Proposition 4.3, $(L(X), <_\ell) = (L(X), <_s)$ is well-ordered, for all $X \in N$.

The following fact is immediate from Lemma 2.2 and the definitions.

Lemma 4.5 If G is a prefix grammar, then for any word $q = v_0 X_1 v_1 \dots v_{k-1} X_k v_k \in (N \cup T)^*$, $L(q)$ is a prefix language with $(L(q), <_\ell) = (L(q), <_s)$ and

$$\mathbf{o}(L(q), <_\ell) = \mathbf{o}(L(X_k), <_\ell) \times \dots \times \mathbf{o}(L(X_1), <_\ell).$$

Moreover, if G is an ordinal grammar, then for any word $q \in (N \cup T)^*$, $(L(q), <_\ell)$ is well-ordered.

We write $\mathbf{o}(q)$ for the order type of $(L(q), <_\ell)$. In particular, $\mathbf{o}(X)$ denotes the order type of the linear order $(L(X), <_\ell) = (L(X), <_s)$.

We list some examples of ordinal grammars. Each grammar includes nonterminals from the set

$$\{\Omega_1, \Omega_2, \dots, \Omega_1^\Omega, \Omega_2^\Omega, \dots\}$$

and productions of some of the previous ones.

1. ω

$$\Omega_1 \rightarrow 0 + 1\Omega_1.$$

$$L(\Omega_1) = 1^*0, \text{ so that } \mathbf{o}(\Omega_1) = \omega.$$

2. ω^2

$$\Omega_2 \rightarrow 0\Omega_1 + 1\Omega_2.$$

$$L(\Omega_2) = \bigcup_{k,n} 1^k 0 1^n 0 = L(\Omega)^2, \text{ so that } \mathbf{o}(\Omega_2) = \omega^2.$$

3. ω^{n+1}

$$\Omega_{n+1} \rightarrow 0\Omega_n + 1\Omega_{n+1}.$$

$$\mathbf{o}(\Omega_{n+1}) = \mathbf{o}(\Omega_1)^{n+1}.$$

4. ω^ω

$$\Omega_1^\Omega \rightarrow 0 + 1\Omega_1^\Omega \Omega_1.$$

$$L(\Omega_1^\Omega) = \bigcup_n 1^n 0 L(\Omega_1)^n = \bigcup_n 1^n 0 (1^* 0)^n, \text{ so that } \mathbf{o}(\Omega_1^\Omega) = \omega^\omega.$$

5. $\omega^{\omega^{n+1}}$

$$\Omega_{n+1}^\Omega \rightarrow 0 + 1\Omega_{n+1}^\Omega \Omega_n^\Omega$$

Note that the first three are regular grammars, but the subsequent ones are context-free ordinal grammars.

We now establish some closure properties of the ordinals of ordinal grammars.

Proposition 4.6 *The set of ordinals $\mathbf{o}(L(G))$, for an ordinal grammar G , is closed under sum, products and ω -powers.*

Proof. Suppose that G_i are ordinal grammars, for $i = 1, 2$. Then, the grammar with a new start symbol S_+ and productions

$$S_+ \rightarrow 0S_1 + 1S_2$$

together with the productions of G_1 and G_2 is a grammar satisfying

$$\mathbf{o}(S_+) = \mathbf{o}(L(G_1)) + \mathbf{o}(L(G_2)).$$

If, instead, we add the new start S_\times and the production

$$S_\times \rightarrow S_2 S_1$$

to the productions of G_1 and G_2 , we obtain a grammar satisfying

$$\mathbf{o}(S_\times) = \mathbf{o}(L(G_1)) \times \mathbf{o}(L(G_2)).$$

Last, if we add the new start symbol S_ω and the productions

$$S_\omega \rightarrow 0 + 1S_\omega S_1$$

to the productions of G_1 , we obtain a grammar satisfying

$$\mathbf{o}(S_\omega) = \mathbf{o}(L(G_1))^\omega.$$

In each case, the constructed grammar is an ordinal grammar. \square

Proposition 4.7 *Any ordinal less than $\omega^{\omega^{\omega}}$ is the ordinal $\mathbf{o}(L(G))$, for some ordinal grammar G .*

Proof. From Proposition 4.6 and Proposition 3.1, using the fact that every finite ordinal is the ordinal of an ordinal grammar. \square

Remark 4.8 *The above proposition also follows from the corresponding result for algebraic ordinals and the facts in Section 5.*

4.2 A bound on ordinals of ordinal grammars

In this section, we will show that for any ordinal grammar G , $\mathbf{o}(L(G))$ is less than $\omega^{\omega^{\omega}}$.

Throughout this section, we assume that G is an ordinal grammar and $L(G)$ is infinite. Moreover, we assume that $L(X)$ contains at least two words, for each nonterminal X . Note that for each nonterminal X , since $L(X)$ is a prefix language, it does not contain the empty word.

Proposition 4.9 *If $X \xRightarrow{*} q$ for some $X \in N$ and $q = v_0 X_1 v_1 \dots v_{k-1} X_k v_k \in (N \cup T)^*$ then $\mathbf{o}(q) = \mathbf{o}(X_k) \times \dots \times \mathbf{o}(X_1) \leq \mathbf{o}(X)$. In particular, $\mathbf{o}(X_i) \leq \mathbf{o}(X)$ for all i .*

Proof. Immediate from Lemma 4.5 and the fact that $L(q) \subseteq L(X)$. \square

The relations \preceq and \approx were defined at the beginning of Section 4.

Corollary 4.10 *If $Y \preceq X$, then $\mathbf{o}(Y) \leq \mathbf{o}(X)$. If $X \approx Y$, $\mathbf{o}(X) = \mathbf{o}(Y)$.*

Proof. Immediate, from Proposition 4.9. \square

It is not necessarily true that if $\mathbf{o}(X) = \mathbf{o}(Y)$ then $X \approx Y$. Consider the example:

$$\begin{aligned} X_1 &\rightarrow 1X_2 \\ X_2 &\rightarrow 1X_2 + 0. \end{aligned}$$

Then $X_2 \prec X_1$, and $\mathbf{o}(X_1) = \mathbf{o}(X_2) = \omega$, since

$$\begin{aligned} L(X_2) &= \bigcup_{n \geq 0} 1^n 0 \\ L(X_1) &= \bigcup_{n \geq 1} 1^n 0. \end{aligned}$$

Proposition 4.11 *For any nonterminal X , there is no derivation $X \xRightarrow{*} Xp$ with $p \neq \epsilon$.*

Proof. Indeed, suppose that $X \xRightarrow{*} Xp$. There are nonempty words u, v such that $p \xRightarrow{*} v$, and $X \xRightarrow{*} u$. Thus, $X \xRightarrow{*} u$ and $X \xRightarrow{*} uv$, so that $L(X)$ is not a prefix language. \square

Proposition 4.12 *For each derivation $X \xRightarrow{*} p$, either $Y \prec X$ holds for all nonterminals Y in p , or there is exactly one nonterminal Y occurring in p with $X \approx Y$, and in this case, $p = uYq$, for some $u \in \{0, 1\}^*$ and for some q with $\text{ht}(q) < \text{ht}(X)$.*

Proof. Suppose that $X \xRightarrow{*} qYrZs$ is a derivation, where $Y, Z \approx X$. Let $\alpha = \mathbf{o}(X)$, $\beta = \mathbf{o}(Y)$, $\gamma = \mathbf{o}(Z)$. Then by Corollary 4.10, $\alpha = \beta = \gamma$, and thus by Proposition 4.9, $\alpha \times \alpha \leq \mathbf{o}(qYrZs) \leq \alpha$. This contradicts the assumption that $L(X)$ contains at least two words. Similarly, if there are nonterminals Y, Z such that $X \xRightarrow{*} qYrZs$ with $\text{ht}(Y) < \text{ht}(X) = \text{ht}(Z)$, then $\mathbf{o}(X) \geq \mathbf{o}(X) \times \mathbf{o}(Z) \geq \mathbf{o}(X) \times 2$, an impossibility, since $\mathbf{o}(X) \geq 2$. \square

The next fact is a basic result.

Lemma 4.13 *Suppose $X \xRightarrow{*} uXp$ and $X \xRightarrow{*} vXq$. If $|u| \leq |v|$, then $u \leq_p v$. In particular, if $|u| = |v|$, then $u = v$.*

Proof. Assume $u <_s v$, say. Suppose $p \xRightarrow{*} w$ and $q \xRightarrow{*} w'$. Then, for each n , $X \xRightarrow{*} u^n X w^n$, so that $X \xRightarrow{*} u^n v z w' w^n = y_n$, where z is any word in $\{0, 1\}^*$ with $X \xRightarrow{*} z$. But, $y_{n+1} <_s y_n$, for each n , contradicting the fact that $(L(X), <_\ell)$ is well-ordered. \square

Lemma 4.14 *Suppose there is a derivation $X \xRightarrow{*} uXp$. Suppose also that $X \xRightarrow{*} v$. Then either $v <_s u$ or $u <_p v$.*

Proof. For any two words u, v there are four possibilities:

$$u <_s v, v \leq_p u, v <_p u, u <_p v.$$

We show that the first two possibilities are ruled out.

If $u <_s v$, then, for any $n \geq 1$, $u^{n+1}vw^{n+1} <_s u^nvw^n$, where w is any terminal word with $p \xRightarrow{*} w$, so that there is a descending chain in $L(X)$.

If $v \leq_p u$, then since $v \neq \epsilon$, $v <_p uvw$, for any word w , so that $L(X)$ is not a prefix language. \square

Recall that a *primitive word* is a nonempty word v which cannot be written as u^n , for any word u and integer $n > 1$. The **primitive root** [Lot97] of a nonempty word v is a primitive word u such that $v = u^n$, for some $n \geq 1$.

Proposition 4.15 *Suppose that X is a recursive nonterminal. Then there is a unique shortest word $u_0^X \in \{0, 1\}^+$ such that whenever $X \xRightarrow{*} uXp$ for some $u \in \{0, 1\}^+$ and $p \in (N \cup T)^*$, then u is a power of u_0^X .*

Proof. Consider any derivation $X \xRightarrow{*} vXq$ with $v \in \{0, 1\}^+$ and $q \in (N \cup T)^*$, let u_0^X denote the primitive root of v . Thus, u_0^X is the shortest word such that v is a power of u_0^X , and clearly, u_0^X is primitive. If $X \xRightarrow{*} uXp$ then there are some $m, n \geq 1$ such that $|u^m| = |v^n|$. But then, by Lemma 4.13,

$$u^m = v^n.$$

It then follows that u, v are powers of the same word (see [Lot97] for example), which implies that u also is a power of u_0^X . \square

Below we will write u_0 for u_0^X whenever X is clear from the context.

Proposition 4.16 *Suppose that X is recursive and $v \in L(X)$. Then, there is some $n \geq 1$ such that $v <_s u_0^n$.*

Proof. Indeed, there is a word u such that $|v| < |u|$ and $X \xRightarrow{*} uXp$, for some $p \in (N \cup T)^*$. Then, $v <_s u$, by Lemma 4.14, and $u = u_0^n$, for some n , by Proposition 4.15. \square

Thus, for $v \in L(X)$, if we choose n as the least integer such that $v <_s u_0^n$, we may write v in a unique way as $u_0^{n-1}w$ where u_0 is not a prefix of w and $w <_s u_0$. Moreover, we can write w as $x0y$ where $x1$ is a prefix of u_0 .

Definition 4.17 *Suppose that X is a recursive nonterminal, and the word $x1$ is a prefix of u_0 . For each $n \geq 0$, define $L(n, x, X)$ as the set of all words of the form $u_0^n x0y$ in $L(X)$. Moreover, define $L(n, X) = \bigcup_x L(n, x, X)$, where x ranges over all words such that $x1$ is a prefix of u_0 .*

Lemma 4.18 *Suppose $n < m$. If $v \in L(n, X)$ and $w \in L(m, X)$, then $v <_s w$.*

Proof. Write $v = u_0^n x0y$ and $w = u_0^m x'0y'$, where $x1$ and $x'1$ are prefixes of u_0 . If $u_0 = x1r$, then

$$\begin{aligned} v &= u_0^n x0y \\ &<_s u_0^n (x1r)^{m-n} x'0y' \\ &= u_0^m x'0y' = w. \quad \square \end{aligned}$$

The following lemma gives an easy upper bound to the ordinal of a union.

Lemma 4.19 *Suppose that L_1, L_2 are subsets of $\{0, 1\}^*$ such that for $i = 1, 2$, $(L_i, <_\ell)$ is well-ordered. Then $(L_1 \cup L_2, <_\ell)$ is well-ordered. Let $\mathbf{o}(L_i, <_\ell) \leq \alpha_i$, where α_i is a (infinite) limit ordinal, so that $1 + \alpha_i = \alpha_i$, $i = 1, 2$. Then*

$$\mathbf{o}(L_1 \cup L_2, <_\ell) \leq \max\{\alpha_1 \times \alpha_2, \alpha_2 \times \alpha_1\}.$$

Proof. To show $L_1 \cup L_2$ is well-ordered, suppose that (v_n) is an infinite descending chain in $L_1 \cup L_2$. Then, either there are infinitely many $v_n \in L_1$, or infinitely many $v_n \in L_2$. Either possibility contradicts the assumption that both L_1 and L_2 are well-ordered.

Without loss of generality we may assume that $(L_1, <_\ell)$ is cofinal in $(L_1 \cup L_2, <_\ell)$, i.e., for each $y \in L_2$ there is some $x \in L_1$ with $y \leq x$. For each $y \in L_2$ let $f(y)$ denote the least $x \in L_1$ with $y \leq x$. For each $x \in L_1$, let $\beta(x)$ denote the order type of the set $\{y \in L_2 : f(y) = x\}$. Then, using Lemma 2.1 in the last line,

$$\begin{aligned} \mathbf{o}(L_1 \cup L_2) &\leq \sum_{x \in L_1} (\beta(x) + 1) \\ &\leq \sum_{x \in L_1} (\alpha_2 + 1) \\ &= (\alpha_2 + 1) \times \alpha_1 \\ &\leq \alpha_2 \times \alpha_1. \end{aligned}$$

Thus, in this case, $\mathbf{o}(L_1 \cup L_2) \leq \alpha_2 \times \alpha_1$. \square

By induction, we have:

Corollary 4.20 *For any finite collection $\{L_i : i = 1, \dots, n\}$ of subsets of $\{0, 1\}^*$ such that for each i , $(L_i, <_\ell)$ is well-ordered, $(\bigcup_i L_i, <_\ell)$ is well-ordered. Moreover, if $\mathbf{o}(L_i, <_\ell) \leq \alpha_i$ where α_i is a limit ordinal for all $1 \leq i \leq n$, it holds that $(\bigcup_i L_i, <_\ell)$*

$$\mathbf{o}\left(\bigcup_i L_i, <_\ell\right) \leq \max_{\pi} \{\alpha_{\pi(1)} \times \dots \times \alpha_{\pi(n)}\}$$

where π ranges over all permutations of $\{1, \dots, n\}$. \square

The next theorem is one of our main results.

Theorem 4.1 *Suppose that X is a nonterminal of height h . Then*

$$\mathbf{o}(X) \leq \omega^{\omega^h}.$$

Proof. We prove this claim by induction on the height of X . Let X be a nonterminal of height h and suppose that we have proved the claim for all

nonterminals of height less than h . Below we will make use of the fact that the set of ordinals less than ω^{ω^h} is closed under sum and product. Moreover, when $h > 0$, then for every ordinal $\alpha < \omega^{\omega^h}$ there is a limit ordinal $\beta < \omega^{\omega^h}$ with $\alpha < \beta$. Indeed, we can choose $\beta = \omega^{h-1} \times n$ for some n .

Case 1. X is not recursive. Then, by the induction hypothesis, we have $\mathbf{o}(Y) < \omega^{\omega^h}$ whenever Y occurs on the right side of a production whose left side is X . It follows by Proposition 4.9 that $\mathbf{o}(p) < \omega^{\omega^h}$ whenever $X \rightarrow p$ is a production. Since $L(X)$ is a finite union of the languages $L(p)$, it follows that $\mathbf{o}(X) < \omega^{\omega^h}$, by Corollary 4.20. (If $h = 0$, $L(p)$ is a single word in $\{0, 1\}^*$. Moreover, $\mathbf{o}(X)$ is finite.)

Case 2. X is recursive. Then for each n, x , $L(n, x, X)$ is a finite union of languages of the form $L(u_0^n x 0 p)$, where there is left derivation $X \xRightarrow{*} u_0^n x 0 p$. It is not possible that p contains a nonterminal Y with $X \approx Y$, since in that case we would have a derivation $X \xRightarrow{*} u_0^n x 0 w X q$ for some terminal word w and some q , contradicting Proposition 4.15. Thus, $Y \prec X$ holds for all nonterminals Y occurring in p . It follows by the induction hypothesis that $\mathbf{o}(L(u_0^n x 0 p), <_\ell) < \omega^{\omega^h}$. Thus, using Corollary 4.20, it follows that $\mathbf{o}(L(n, x, X), <_\ell) < \omega^{\omega^h}$. Again, by Corollary 4.20,

$$\begin{aligned} \alpha_n &= \mathbf{o}(L(n, X), <_\ell) \\ &< \omega^{\omega^h}. \end{aligned}$$

(If $h = 0$, $\alpha_n < \omega$, for all n .) But, by Lemma 4.18,

$$\begin{aligned} \mathbf{o}(X) &= \alpha_0 + \alpha_1 + \dots \\ &= \sup\left\{\sum_{i \in [n]} \alpha_i : n \geq 0\right\}. \end{aligned}$$

Since $\sum_{i \in [n]} \alpha_i < \omega^{\omega^h}$ for all n , it follows that $\mathbf{o}(X) \leq \omega^{\omega^h}$. \square

Corollary 4.21 *If G is an ordinal grammar, there is an integer n such that*

$$\mathbf{o}(G) \leq \omega^{\omega^n}.$$

Proof. Let n be the height of the start symbol S . Then, by Theorem 4.1, $\mathbf{o}(S) \leq \omega^{\omega^n}$. \square

We have thus completed the proof of this characterization of the ordinals of ordinal grammars:

Theorem 4.2 *An ordinal α is less than ω^{ω^ω} if and only if there is some ordinal grammar G such that $\alpha = \mathbf{o}(L(G), <_\ell)$.*

5 From algebraic trees to prefix grammars

In this section we show that each system of equations defining an algebraic tree can be transformed (in polynomial time) to a prefix grammar⁴ generating the frontier of the tree. This result allows us to complete the proof of the fact that every algebraic ordinal is less than $\omega^{\omega^{\omega}}$.

Consider a system of equations

$$F_i(x_0, \dots, x_{n_i-1}) = t_i(x_0, \dots, x_{n_i-1}), \quad i = 1, \dots, m, \quad (6)$$

where each t_i is a term over the ranked alphabet $\Sigma \cup \mathcal{F}$ in the variables x_0, \dots, x_{n_i-1} . We assume that F_1 is the principal function variable and that $n_1 = 0$. Each component of the least solution is an algebraic tree. Let (T_1, T_2, \dots, T_m) denote the least solution of the system (6).

For the ranked alphabet $\Sigma \cup \mathcal{F}$, let Δ be the (unranked) alphabet whose letters are the letters $(\sigma, k), (F_i, j)$ where $\sigma \in \Sigma_n$, $n > 0$ and $k \in [n]$, $j \in [n_i]$ and $1 \leq i \leq m$. Let T be a finite or infinite tree in $T_{\Sigma \cup \mathcal{F}}^{\omega}(V)$.

Definition 5.1 *For each vertex $u \in \text{dom}(T)$ we define a word $\hat{u} \in \Delta^*$ by induction. First, $\hat{\epsilon} = \epsilon$. When $u = vi$ and $T(v) = \delta$, then $\delta \in (\Sigma \cup \mathcal{F})_k$, for some $k > 0$, and we define $\hat{u} = \hat{v}(\delta, i)$.*

We define the “labeled frontier language of $T \in T_{\Sigma}^{\omega}$ ” as the set of words

$$\text{Lfr}(T) = \{\hat{u}T(u) : T(u) \in \Sigma_0\}$$

We prove the following.

Theorem 5.1 *When $T \in T_{\Sigma}^{\omega}$ is an algebraic tree, $\text{Lfr}(T)$ can be generated by a prefix grammar.*⁵

Proof. Suppose that T is the principal component of the least solution of the system (6).

We will define a grammar whose nonterminals N consist of the letters F_i , together with all ordered pairs (F_i, j) where $i = 1, \dots, m$, $j \in [n_i]$.

The grammar will be designed to have the following properties.

Claim: *For any word u , $T_i(u) = x_j$ if and only if $(F_i, j) \xrightarrow{*} \hat{u}$. And for any word u , $T_i(u) \in \Sigma_0$ if and only if $F_i \xrightarrow{*} \hat{u}T_i(u)$. Moreover, any terminal word*

⁴Here, we allow grammars over an arbitrary (linearly ordered) terminal alphabet. The notion of a prefix grammar can be adjusted appropriately.

⁵In [Cour83], $\text{Lfr}(T)$ is called the *branch language* of T . Courcelle showed that a “locally finite” tree T is algebraic if and only if $\text{Lfr}(T)$ is a strict deterministic context-free language, see [Cour78b, Cour83].

derivable from (F_i, j) is of the form \hat{u} , and any terminal word derivable from F_i is of the form $\hat{u}T_i(u)$ for some $u \in \text{dom}(T_i)$.

Let $\Gamma = \Sigma_0 \cup \{(\sigma, j) : \sigma \in \Sigma_k, j \in [k]\}$. The grammar generating $\text{Lfr}(T)$ is: $G_L = (N, \Gamma, P, F_1)$, where $N = \mathcal{F} \cup \{(F_i, j) : 1 \leq i \leq m, j \in [n_i]\}$ and the set P of productions is defined below. If t_1, \dots, t_m are the terms on the right side of (6) above, then the productions are:

- $$(F_i, j) \rightarrow \hat{u}$$

where $u \in \text{dom}(t_i)$ and $t_i(u) = x_j$,
- $$F_i \rightarrow \hat{u}t_i(u)$$

where $u \in \text{dom}(t_i)$ and $t_i(u) \in \Sigma_0 \cup \mathcal{F}$.

The proof of the fact that the above grammar is a prefix grammar generating the language $\text{Lfr}(T)$ relies on the above claim, and may be found in the Appendix. \square

Example. Suppose the system of equations is:

$$\begin{aligned} F_0 &= F_1(a) \\ F_1(x) &= F_2(a, x) \\ F_2(x, y) &= \sigma(x, a, F_2(x, F_2(x, y))) \end{aligned}$$

where the individual variables x, y stand for x_0, x_1 , respectively. Then the productions in G_L are:

$$\begin{aligned} F_0 &\rightarrow F_1 + (F_1, 0)a \\ F_1 &\rightarrow F_2 + (F_2, 0)a \\ (F_1, 0) &\rightarrow (F_2, 1) \\ (F_2, 0) &\rightarrow (\sigma, 0) + (\sigma, 2)(F_2, 0) + (\sigma, 2)(F_2, 1)(F_2, 0) \\ (F_2, 1) &\rightarrow (\sigma, 2)(F_2, 1)(F_2, 1) \\ F_2 &\rightarrow (\sigma, 1)a + (\sigma, 2)F_2 + (\sigma, 2)(F_2, 1)F_2 \end{aligned}$$

Corollary 5.2 *For every system of equations defining an algebraic tree $T \in T_\Sigma^\omega$ one can construct in polynomial time a prefix grammar generating the frontier of T .*

Proof. To get from the prefix grammar G_L which derives $\text{Lfr}(T)$ to a grammar G' which derives $\text{Fr}(T)$, replace each letter (σ, j) by just j , and delete the

constant symbols in Σ_0 . Thus, in the example above, the productions of G' are

$$\begin{aligned} F_0 &\rightarrow F_1 + (F_1, 0) \\ F_1 &\rightarrow F_2 + (F_2, 0) \\ (F_1, 0) &\rightarrow (F_2, 1) \\ (F_2, 0) &\rightarrow 0 + 2(F_2, 0) + 2(F_2, 1)(F_2, 0) \\ (F_2, 1) &\rightarrow 2(F_2, 1)(F_2, 1) \\ F_2 &\rightarrow 1 + 2F_2 + 2(F_2, 1)F_2. \end{aligned}$$

It follows that G' is a prefix grammar generating $\text{Fr}(T)$. \square

Corollary 5.3 *If α is an algebraic ordinal, there is an ordinal grammar G' with $\alpha = \mathbf{o}(L(G'))$.*

We may now derive our main theorem.

Theorem 5.2 *An ordinal is algebraic if and only if it is less than ω^{ω^ω} .*

We needed only to prove the “only if” direction. But this follows immediately from Corollary 5.3 and Theorem 4.2. \square

6 Conclusion

We have proved that the algebraic ordinals are exactly those less than ω^{ω^ω} , or equivalently, the ordinals that can be constructed from 0 and 1 by the sum and product operations, and the operation $\alpha \mapsto \alpha^\omega$. It is known that the regular ordinals are those less than ω^ω , or equivalently, those that can be constructed from 0, 1 and ω by just sum and product; or the ordinals that can be constructed from 0 and 1 by sum, product, and the operation $\alpha \mapsto \alpha \times \omega$.

Recall (from [Ros82], for example) that the **Hausdorff rank** of a countable *scattered* linear ordering L is the least ordinal α such that $L \in V_\alpha$, where

$$V_0 := \{0, 1\},$$

and if $\alpha > 0$, the collection V_α is defined by:

$$V_\alpha = \left\{ \sum_{i \in I} L_i : L_i \in \bigcup_{\beta < \alpha} V_\beta \right\},$$

where I is either ω , or ω^* , the reverse of ω , or a finite ordinal n , or $\omega^* + \omega$.

In order to characterize the algebraic *linear orders*, the next step might be the characterization of algebraic *scattered* linear orders. We conjecture that

any such linear order has Hausdorff rank less than ω^ω . Moreover, one possible conjecture is that, up to isomorphism, these are the linear orders that can be constructed from the empty linear order and a one point linear order by the sum and product operations, reversal, and the operation $P \mapsto P^\omega$, where

$$P^\omega = \sum_{n \in \omega} P^n.$$

This conjecture is supported by the fact that the scattered regular linear orders are exactly those that can be constructed from 0, 1 and ω by the sum and product operations and reversal, cf. [Heil80]. Thus, the scattered regular linear orders have finite Hausdorff rank, but the converse is false: the linear order

$$\mathbb{Z} + 1 + \mathbb{Z} + 2 + \mathbb{Z} + \dots + \mathbb{Z} + n + \mathbb{Z} + \dots$$

where \mathbb{Z} denotes the linear order of the negative and positive integers is not regular, but has finite Hausdorff rank.

After describing the scattered algebraic linear orders, the next task could be to obtain a characterization of all algebraic linear orders. We conjecture that these are exactly those linear orders that can be constructed from dense algebraic words by substituting a scattered algebraic linear order for each letter. (See below for the definition of an algebraic word.) Thus, the task can be reduced to the characterization of the dense algebraic words.

A hierarchy of recursion schemes was studied by [Damm77, Damm82, Gal84, Ong07, HMOS07], and many others. The schemes considered in this paper are on the first level of the hierarchy with regular schemes forming level 0. In the light of the characterizations of the regular and algebraic ordinals, it is natural to conjecture that the ordinals definable on the n th level of the hierarchy are those less than

$$\uparrow (\omega, n + 2) = \omega^{\omega^{\vdots^\omega}}$$

where there are $n + 2$ ω 's altogether. In fact, every ordinal less than $\uparrow (\omega, n + 2)$ is shown to be definable on the n th level in [Braud].

In ordinal analysis of logical theories, the strength of a theory is measured by ordinals. For example, the proof theoretic ordinal of Peano arithmetic is ϵ_0 . Here we have a similar phenomenon: we measure the strength of recursive definitions by ordinals, and we conjecture that the definable ordinals are exactly those less than ϵ_0 .

A generalization of the notion of “finite word” is obtained by considering labeled linear orders, where the labels are letters in some finite alphabet. Thus, a linear order may be identified with a word on a one letter alphabet. A **countable word** is word whose underlying linear order is countable. A morphism between words is a morphism between their respective underlying linear orders that additionally preserves the labeling. Every countable word can be represented as

the word determined by the frontier of a tree where each leaf retains its label, cf. [Cour78a]. Now an **algebraic word** (respectively **regular word**) is a word isomorphic to the frontier word of an algebraic (respectively, regular) tree. An “operational” characterization of the regular words was obtained in [Heil80], where it was shown that a nonempty word is regular if and only if it can be constructed from single letter words by concatenation, ω -power, the “shuffle operations” and reversal. (Note that concatenation corresponds to the sum operation on linear orders, and ω -power to the operation $P \mapsto P \times \omega$.) Without the shuffle operations, exactly the nonempty scattered regular words can be generated, and the well-ordered regular words can be generated by concatenation and ω -power. It would be interesting to obtain operational characterizations of well-ordered, scattered, and eventually, all algebraic words.

Finally, we would like to mention an open problem. Suppose that a context-free language L is well-ordered by the lexicographic order. Is the order type of $(L, <_\ell)$ less than ω^{ω^ω} ?

7 Acknowledgement

The authors would like to thank the three referees whose suggestions have resulted in an improved paper.

References

- [BC01] S.L. Bloom and C. Choffrut. Long words: the theory of concatenation and ω -power. *Theoretical Computer Science*, 259(2001), 533–548.
- [BE04] S.L. Bloom and Z. Ésik. Deciding whether the frontier of a regular tree is scattered. *Fundamenta Informaticae*, XI(2004) 1-22.
- [BE07] S.L. Bloom and Z. Ésik. Regular and algebraic words and ordinals. *CALCO 2007, Bergen*, LNCS 4624, Springer, 2007, 1–15.
- [BE10] S. L. Bloom and Z. Ésik. A Mezei-Wright theorem for categorical algebras. *Theoretical Computer Science*, 411(2010), 341–359
- [Braud] L. Braud. Unpublished paper. ENS-Lyon, MR2 Intern at IRISA, Rennes.
- [BKN08] M. Brough, B. Khoussainov and P. Nelson. Sequential automatic algebras. *CiE 2008*, LNCS 5028, Springer, 2008, 84–93.
- [Col04] Th. Colcombet. Equational presentations of tree automatic structures. *Workshop on Automata, Structures, and Logic*, Auckland, 2004.

- [Cour78a] B. Courcelle. Frontiers of infinite trees. *Theoretical Informatics and Applications*, 12(1978), 319–337.
- [Cour78b] B. Courcelle. A representation of trees by languages, *Theoretical Computer Science*, 6(1978), 255–279 and 7(1978), 25–55.
- [Cour83] B. Courcelle. Fundamental properties of infinite trees. *Theoretical Computer Science*, 25(1983), 95–169.
- [Damm77] W. Damm. Higher type program schemes and their tree languages. *Theoretical Computer Science (Third GI Conf., Darmstadt, 1977)*, LNCS 48, Springer, 1977, 51–72.
- [Damm82] W. Damm. The IO and OI hierarchies. *Theoretical Computer Science*, 20(1982), 95–206.
- [DT90] M. Dauchet and S. Tison. The theory of ground rewrite systems is decidable. *LICS 1990*, IEEE, 1990, 242–248
- [Del04] Ch. Delhommé. Automaticity of ordinals and of homogeneous graphs. *C. R. Math. Acad. Sci. Paris*, 339(2004), 5–10 (in French).
- [Gal84] J. Gallier. n -rational algebras. I. Basic properties and free algebras. *SIAM J. Comput.*, 13(1984), 750–775.
- [GTWW77] J. A. Goguen, J. W. Thatcher, E. G. Wagner and J. B. Wright. Initial algebraic semantics and continuous algebras. *J. ACM*, 24(1977), 68–95.
- [Gue81] I. Guessarian. *Algebraic Semantics*. LNCS, vol. 99, Springer-Verlag, 1981.
- [GS84] F. Gécseg and M. Steinby. *Tree Automata*. Akadémiai Kiadó, Budapest, 1984.
- [Heil80] S. Heilbrunner. An algorithm for the solution of fixed-point equations for infinite words. *Theoretical Informatics and Applications*, 14(1980), 131–141.
- [Hod82] B.R. Hodgson. On direct products of automaton decidable theories. *Theoretical Computer Science*, 19(1982), 331–335.
- [HMOS07] M. Hague, A.S. Murawawski, C.-H Luke Ong, O. Serre. Collapsible pushdown automata and recursion schemes. *23rd Annual IEEE Symposium on Logic in Computer Science, LICS*, IEEE, 2008, 452–461.
- [KN95] B. Khoussainov, A. Nerode. Automatic presentations of structures. *Logic and Comput. Complex.*, LNCS 960, Springer, 1995, 367–392.
- [KRS03] B. Khoussainov, S. Rubin and F. Stephan. On automatic partial orders. *Eighteenth IEEE Symposium on Logic in Computer Science, LICS*, IEEE 2003, 168–177.

- [Ong07] C.-H Luke Ong. Hierarchies of infinite structures generated by push-down automata and recursion schemes. *MFCS 2007*, LNCS 4708, Springer, 2007, 15–21.
- [Lot97] M. Lothaire. *Combinatorics on Words*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1997.
- [MW67] J. Mezei and J. Wright. Algebraic automata and context-free sets. *Information and Control*, 11(1967), 3–29.
- [Roit90] J. Roitman. *Introduction to Modern Set Theory*. Wiley, 1990.
- [Ros82] J.B. Rosenstein. *Linear Orderings*. Academic Press, New York, 1982.
- [Sier58] W. Sierpinski. *Cardinal and Ordinal Numbers*. Warsaw: PWN, 1958.
- [Thom86] W. Thomas. On frontiers of regular trees. *Theoretical Informatics and Applications*, vol. 20, 1986, 371–381.

Appendix

This appendix is devoted to a formal proof of the correctness of the translation given in Section 5.

Consider the system (6) whose least solution in $T_\Sigma^\omega(V)$ is (T_1, T_2, \dots, T_m) . To this system, we can associate the *tree grammar* G_T whose productions are

$$F_i(x_0, \dots, x_{n_i-1}) \rightarrow t_i + \perp, \quad i = 1, \dots, m,$$

where \perp denotes the empty tree. The start symbol is F_1 . Let V denote the set of individual variables that occur in (6).

Below we will also assume a new individual variable z and write $t = t' \star t''$ for a tree t over $\Sigma \cup \mathcal{F}$ possibly containing variables in V if t' is a tree with a *single* leaf labeled z and t is obtained from t' by replacing this leaf with a copy of t'' .

Definition 7.1 *Suppose that t, t' are finite trees in $T_\Sigma^\omega(V)$. Then $t \Rightarrow t'$ in G_T if t can be written as $s \star F_i(s_0, \dots, s_{n_i-1})$ for some trees s and s_0, \dots, s_{n_i-1} such that $t' = s \star t_i(s_0, \dots, s_{n_i-1})$ or $t' = s \star \perp$. The relation \Rightarrow^* is the reflexive transitive closure of \Rightarrow .*

It is known, cf., [Cour83, Gue81], that for any $i = 1, \dots, m$ and for any word u , $T_i(u)$ is defined if and only if there is some finite tree t in $T_\Sigma^\omega(V)$ with $F_i(x_0, \dots, x_{n_i-1}) \Rightarrow^* t$ such that $t(u)$ is defined, and in that case $T_i(u) = t(u)$.

Thus, it suffices to prove that the grammar G_L defined in Section 5 and the tree grammar G_T are related as follows:

Lemma 7.2 Let $i \in \{1, \dots, m\}$ and let t be a finite tree $T_\Sigma^\omega(V)$. Suppose that $F_i(x_0, \dots, x_{n_i-1}) \xRightarrow{*} t$. Then for every $u \in \text{dom}(t)$ and $j \in [n_i]$, if $t(u) = x_j$ then $(F_i, j) \xRightarrow{*} \hat{u}$, and if $t(u) \in \Sigma_0 \cup \mathcal{F}$ then $F_i \xRightarrow{*} \hat{u}t(u)$.

Lemma 7.3 Let $i \in \{1, \dots, m\}$ and $j \in [n_i]$.

1. Suppose that $(F_i, j) \xRightarrow{*} w$. Then there exist a finite tree t in $T_\Sigma^\omega(V)$ and a word $u \in \text{dom}(t)$ such that $t(u) = x_j$, $w = \hat{u}$ and $F_i(x_0, \dots, x_{n_i-1}) \xRightarrow{*} t$.
2. Suppose that $F_i \xRightarrow{*} w$. Then there exist t and u as above with $t(u) \in \Sigma_0 \cup \mathcal{F}$, $w = \hat{u}t(u)$ and $F_i(x_0, \dots, x_{n_i-1}) \xRightarrow{*} t$.

Proof of Lemma 7.2. We argue by induction on the length of the derivation. When the length is 0, $t = F_i(x_0, \dots, x_{n_i-1})$. If $u = j$, for some $j \in [n_i]$, then $\hat{u} = (F_i, j)$ and we clearly have $(F_i, j) \xRightarrow{*} (F_i, j) = \hat{u}$. If $u = \epsilon$ then $\hat{u} = \epsilon$, $t(\epsilon) = F_i$ and we have $F_i \xRightarrow{*} F_i = \hat{u}t(u)$.

In the induction step, assume that the length of the derivation is positive and that our claim holds for all derivations of smaller length. Suppose that $t(u)$ is a variable x_j or $t(u) \in \Sigma_0 \cup \mathcal{F}$. Let us write the derivation as

$$F_i(x_0, \dots, x_{n_i-1}) \xRightarrow{*} t' \Longrightarrow t$$

where in the last step we have $t' = s \star F_k(s_0, \dots, s_{n_k-1})$ and $t = s \star t_k(s_0, \dots, s_{n_k-1})$ or $t = s \star \perp$. In the second case, $u \in \text{dom}(t')$, moreover \hat{u} in t' is the same as \hat{u} in t , or as \hat{u} in s . Moreover, $t(u) = s(u) = t'(u)$. Thus, using the induction hypothesis, we obtain $(F_i, j) \xRightarrow{*} \hat{u}$ or $F_i \xRightarrow{*} \hat{u}t'(u) = \hat{u}t(u)$ according to whether $t(u) = x_j$ for some $j \in [n_i]$ or $t(u) \in \Sigma_0 \cup \mathcal{F}$.

Assume now that $t' = s \star F_k(s_0, \dots, s_{n_k-1})$ and $t = s \star t_k(s_0, \dots, s_{n_k-1})$. Let v_0 denote the unique word with $s(v_0) = z$. There are two cases. If v_0 is not a prefix of u , then we have that $u \in \text{dom}(t') \cap \text{dom}(s)$, \hat{u} in t is the same as \hat{u} in t' , and $t(u) = s(u) = t'(u)$. The proof is completed as before. So let v_0 be a prefix of u . If there is some $w \in \text{dom}(t_k)$ such that $u = v_0w$ and $t_k(w)$ is not an individual variable, then \hat{u} in t is $\hat{v}_0\hat{w}$, where \hat{v}_0 is computed in s and \hat{w} is computed in t_k . Moreover, $t(u) = t_k(w) \in \Sigma_0 \cup \mathcal{F}$. By the induction hypothesis we have $F_i \xRightarrow{*} \hat{v}_0F_k$, and by construction, $F_k \rightarrow \hat{w}t_k(w)$ is a production. Thus, $F_i \xRightarrow{*} \hat{v}_0\hat{w}t'(v_0w) = \hat{v}_0\hat{w}t_k(w) = \hat{u}t(u)$.

Suppose last that $u = v_0wv_1$, where $w \in \text{dom}(t_k)$ with $t_k(w) = x_h$ for some $h \in [n_k]$ and $v_1 \in \text{dom}(s_h)$. In that case \hat{u} in t is $\hat{v}_0\hat{w}\hat{v}_1$, where \hat{v}_0 and \hat{w} are as before, and \hat{v}_1 is computed in s_h . Moreover, $t(u) = s_h(v_1) = t'(v_0hv_1)$. Assume that $t(u)$ is the individual variable x_j . Then $(F_i, j) \xRightarrow{*} \hat{v}_0(F_k, h)\hat{v}_1$ by the induction hypothesis, moreover, $(F_k, h) \rightarrow \hat{w}$ is a production. We conclude that $(F_i, j) \xRightarrow{*} \hat{v}_0\hat{w}\hat{v}_1 = \hat{u}$. Suppose now that $t(u) \in \Sigma_0 \cup \mathcal{F}$. Then $F_i \xRightarrow{*} \hat{v}_0(F_k, h)\hat{v}_1t'(v_0hv_1) = \hat{v}_0(F_k, h)\hat{v}_1s_h(v_1) \Longrightarrow \hat{v}_0\hat{w}\hat{v}_1s_h(v_1) = \hat{u}t(u)$. \square

Proof of Lemma 7.3. Suppose first that $(F_i, j) \xRightarrow{*} w$. If the length of the derivation is 0, our claim is trivial: let $t = F_i(x_0, \dots, x_{n_i-1})$, $u = j$. We proceed by induction. In the induction step, we can write the derivation as $(F_i, j) \xRightarrow{*} w_0(F_k, h)w_1 \Rightarrow w_0qw_1 = w$, where, by the induction hypothesis, there exist some t' and u_0, u_1 with $t'(u_0hu_1) = x_j$, $w_0(F_k, h)w_1 = \widehat{u_0hu_1}$ in t' and $F_i(x_0, \dots, x_{n_i-1}) \xRightarrow{*} t'$. Since $(F_k, h) \rightarrow q$ is a production of G_L , there is some p with $t_k(p) = x_h$ and $\widehat{p} = q$ in t_k . Clearly, we can write t' as $t' = s \star F_k(s_0, \dots, s_{n_k-1})$ where $s(u_0) = z$, so that $u_1 \in \text{dom}(s_h)$ with $s_h(u_1) = x_j$. Now let $t = s \star t_k(s_0, \dots, s_{n_k-1})$ and consider the word $u = u_0pu_1$. We have that $t(u) = s_h(u_1) = t'(u_0hu_1) = x_j$ and $\widehat{u} = \widehat{u_0p}u_1 = w_0qw_1 = w$ in t .

Suppose next that $F_i \xRightarrow{*} w$. If the length of the derivation is 0, then $w = F_i$ and we take $t = F_i(x_0, \dots, x_{n_i-1})$ and $u = \epsilon$. Assume now that the length of the derivation is positive and that our claim holds for shorter derivations. We can decompose the derivation either as

$$F_i \xRightarrow{*} w_0(F_k, h)w_1 \Rightarrow w_0qw_1 = w$$

or as

$$F_i \xRightarrow{*} w_0F_k \xRightarrow{*} w_0w_1 = w.$$

The former case is similar to the previous one, so we only deal with the latter. In this case, by the induction hypothesis, there is some tree t' and a word $u_0 \in \text{dom}(t')$ with $t'(u_0) = F_k$, $w = \widehat{u_0}t'(u_0)$ and $F_i(x_0, \dots, x_{n_i-1}) \xRightarrow{*} t'$. Moreover, $F_k \rightarrow w_1$ is a production and thus $w_1 = \widehat{u_1}t_k(u_1)$ in t_k for some u_1 with $t_k(u_1) \in \Sigma_0 \cup \mathcal{F}$. Since $t'(u_0) = F_k$, we can write t' as $t' = s \star F_k(s_0, \dots, s_{n_k-1})$, where $s(u_0) = z$. Now let $t = s \star t_k(s_0, \dots, s_{n_k-1})$ and let $u = u_0u_1$. We have that $F_i(x_0, \dots, x_{n_i-1}) \xRightarrow{*} t$ and $w = w_0w_1 = \widehat{u}t(u)$. Moreover, $t(u) = t_k(u_1) \in \Sigma_0 \cup \mathcal{F}$. \square

We can now prove the claim formulated in the proof of Theorem 5.1. Suppose that $T_i(u) = x_j$ for some i , $1 \leq i \leq m$ and $j \in [n_i]$. Then $(F_i, j) \xRightarrow{*} \widehat{u}$ by Lemma 7.2. And if $T_i(u) \in \Sigma_0$, then $(F_i, j) \xRightarrow{*} \widehat{u}T_i(u)$. Conversely, if $(F_i, j) \xRightarrow{*} w$ for some terminal word w , then by Lemma 7.3, either $w = \widehat{u}$ for some u with $T_i(u) = x_j$, or $w = \widehat{u}T_i(u)$ for some u with $T_i(u) \in \Sigma_0$. \square